



PROGRAMA INSTITUCIONAL DE BOLSAS DE INICIAÇÃO
CIENTÍFICA - PIBIC

RELATÓRIO PARCIAL

**Predição de Séries Temporais Baseada em Redes Neurais
Artificiais**

Submetido à
Pró-Reitoria de Pesquisa da Universidade Estadual de Campinas

Departamento de Engenharia de Computação e Automação Industrial (DCA)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas (UNICAMP)
CEP 13083-852, Campinas - SP

Candidato: João Pedro de Oliveira Pagnan
Orientador: Prof. Levy Boccato

Campinas, 26 de fevereiro de 2021

1 Introdução

A predição de séries temporais é uma das aplicações mais interessantes do tratamento de informação. A tarefa de antecipar padrões de comportamento e construir modelos que sejam apropriados para explicar determinados fenômenos da natureza tem importância para a biologia, economia, automação industrial, meteorologia e diversas outras áreas da ciência [1].

Diversos tipos de preditores existem para a predição de séries temporais, desde métodos clássicos lineares como o modelo autoregressivo (AR) até métodos não-lineares utilizando, por exemplo, redes neurais artificiais, sendo que dessas se destacam as redes do tipo *Multilayer Perceptron* (MLP) e as redes recorrentes, especialmente a *Long Short-Term Memory* (LSTM) [2].

Um tipo de sistema que pode ser descrito através de séries temporais é um sistema com dinâmica caótica. Diversos fenômenos naturais, como a dinâmica populacional de uma espécie, a dinâmica atmosférica de uma região, ou até mesmo as órbitas de um sistema com três ou mais corpos celestes são caóticos. Esses sistemas se destacam por, apesar de serem determinísticos (e, portanto, previsíveis), são extremamente sensíveis às condições iniciais [3]. Isso causa um problema para a predição das séries temporais desses sistemas, afinal, uma pequena incerteza na medida afetará toda a previsão.

Tendo em vista o desempenho de modelos não-lineares para previsão de diversas séries temporais [2], optamos por estudar a aplicabilidade de modelos não-lineares utilizando redes neurais artificiais na predição de séries temporais de sistemas com dinâmica caótica.

Essa primeira parte da atividade de iniciação científica teve como objetivo estudar a base teórica das redes neurais artificiais e de outros regressores lineares clássicos, assim como estudar sistemas com dinâmica caótica. Os principais modelos estudados e a caracterização dos sistemas caóticos são brevemente apresentados na Seção 2.

O estudo de redes neurais e regressores foi conduzido não só através da pesquisa da iniciação científica como também através de uma disciplina de pós-graduação da Faculdade de Engenharia Elétrica e Computação da Universidade Estadual de Campinas sobre aprendizado de máquina, cursada como estudante especial. Nela, foram desenvolvidos alguns ensaios de aplicações dessas ferramentas, que são apresentados na Seção 3.

Por fim, na Seção 4 são indicados os próximos passos e metas dessa atividade, concluindo essa primeira metade da iniciação científica.

2 Redes Neurais e Sistemas Caóticos

2.1 Redes Neurais Artificiais

As redes neurais **artificiais** são ferramentas computacionais cujas estruturas são inspiradas no funcionamento das redes neurais biológicas presentes em cérebros **animais**. Podemos interpretar um neurônio (tanto biológico, quanto artificial) como uma unidade de processamento de informação [4]. Os neurônios **artificiais** podem ser organizados de diversas formas, assim como **serem estruturados de forma diferente dependendo da arquitetura desejada para a aplicação**.

Veremos então os dois principais modelos de redes que serão utilizados nessa pesquisa:

2.1.1 Redes *Multilayer Perceptron* (MLP)

Um dos modelos mais utilizados para representar um neurônio artificial, o *Perceptron* [5], é apresentado na Figura 1.

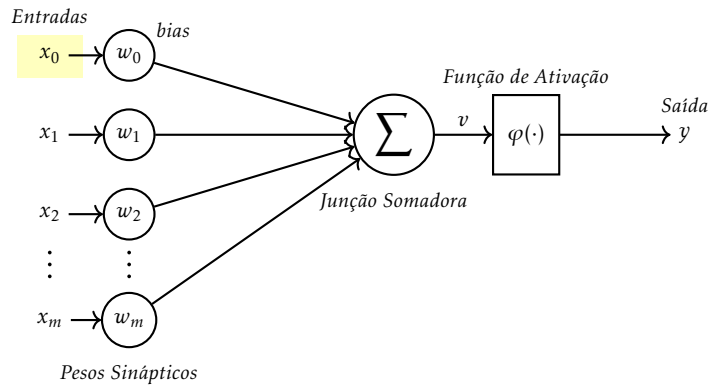


Figura 1: Modelo *Perceptron* para o neurônio artificial

~~Em especial, temos que $x_0 = 1$, dessa forma, o termo w_0 é conhecido como *bias*.~~ Logo, em termos matemáticos:

$$y = \varphi(v) = \varphi\left(\sum_{i=1}^m w_i x_i + w_0\right) = \varphi\left(\sum_{i=0}^m w_i x_i\right) = \varphi(\mathbf{w}^T \cdot \mathbf{x}) \quad (1)$$

A escolha da função de ativação $\varphi(\cdot)$ varia de acordo com a aplicação desejada. Ela pode ser desde uma função de *Heaviside*, a puramente linear $\varphi(x) = x$, ou até mesmo a tangente hiperbólica, a função logística ou outras funções não-lineares para mapeamentos mais complexos [6]. Os pesos sinápticos w são ajustados com um processo iterativo de forma a minimizar uma função custo $J(\mathbf{w})$ que representa uma medida do erro entre as saídas geradas pela rede e as saídas desejadas.

Tipicamente, uma rede neural MLP é composta por um número arbitrário k de camadas com n neurônios do tipo *Perceptron*.

O grande apelo das redes *Multilayer Perceptron* é que elas tem a capacidade de aproximação universal, ou seja, são capazes de aproximar qualquer mapeamento contínuo num domínio compacto com um nível de erro arbitrariamente pequeno. Até mesmo uma MLP com uma única camada intermediária e camada de saída linear já consegue realizar esse mapeamento [7].

2.1.2 Redes Recorrentes *Long Short-term Memory* (LSTM)

Diferentemente das redes MLP que são *feedforward*, ou seja, que não reutilizam a informação processada dos padrões anteriores para gerar a próxima saída, a ideia central das redes recorrentes é que elas têm estruturas computacionais que podem armazenar os estados anteriores dos neurônios, possuindo também portas não-lineares que regulam o fluxo de informação de entrada e de saída da célula computacional [4]. Uma representação possível de uma célula de uma rede recorrente pode ser vista na Figura 2.

Em especial, as redes LSTM são capazes de criar 'memória' de curto e longo prazo. A estrutura dela é apresentada na Figura 3. As LSTMs manipulam o vetor $c(t)$, aprendendo durante o treinamento o que deve ser guardado nele, o que deve ser descartado e o que deve ser aproveitado para gerar a saída $h(t)$.

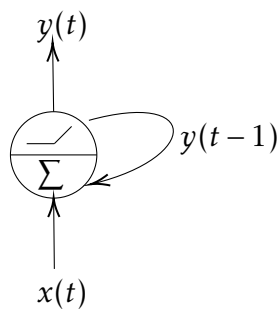


Figura 2: Célula da rede recorrente

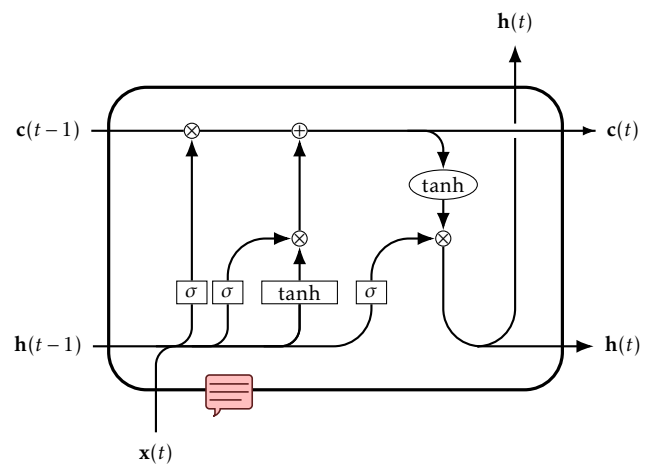


Figura 3: Estrutura de uma célula/camada LSTM

2.2 Sistemas com Dinâmica Caótica

Como dito anteriormente, sistemas com dinâmica caótica se destacam pois, apesar de serem determinísticos, apresentam dependência sensível das condições iniciais (DSCI). Dessa forma, duas trajetórias **que** partem de posições relativamente próximas podem evoluir de uma forma totalmente distinta devido às não-linearidades presentes que amplificam as diferenças entre essas condições iniciais [3].

Podemos caracterizar a dinâmica caótica através das seguintes aspectos [8]:

1. **Um sistema caótico possui DSCI;**
2. A evolução temporal das variáveis de estado (parâmetros **da** ordem do sistema) é rápida e tem uma aparência errática;
3. Tem **espectro de potências contínuos e de banda larga;**
4. Há uma produção de informação por parte do sistema;
5. Dão origem a atratores estranhos (estruturas topológicas que ditam a evolução temporal do fluxo de um sistema caótico) [9].

3 Ensaios de aplicações de Aprendizado de Máquina

Como já foi mencionado, para complementar os estudos **de** iniciação científica, **foi cursada** como estudante especial uma disciplina de pós-graduação sobre aprendizado de máquina da FEEC. O objetivo foi formar uma base sólida para o uso não só de redes neurais MLP, LSTM e regressores, como também noções de probabilidade, teoria da informação, classificadores, árvores de decisão, clusterizadores e outras ferramentas de *Machine Learning*.

Algumas das atividades notórias desenvolvidas na disciplina foram o uso de modelos clássicos de regressores lineares e não-lineares para a predição de uma série temporal do número de **mancha** solares observado, testes com classificadores utilizando redes neurais MLP **e também convolucionais para cenários binários e multiclasse**, e o desenvolvimento de um projeto final composto por um *Autoencoder* utilizando redes convolucionais e profundas para a filtragem de sinais.

Com isso, além de reforçar a base teórica necessária para essa pesquisa, foi obtida uma prática com a programação, análise, teste, otimização e implementação de algoritmos de redes neurais **para diversos cenários**.

4 Próximos Passos

Como nessa primeira parte da iniciação o foco foi uma pesquisa bibliográfica dos temas a serem estudados nela, a segunda metade será voltada para a aplicação em si da

predição das séries temporais de sistemas caóticos.

Primeiramente, escolheremos as séries que serão analisadas, que incluem, mas não estão limitadas a, mapas logísticos, dados meteorológicos, dados da dinâmica populacional de uma espécie etc.

Em seguida, determinaremos aspectos mais fundamentais das redes neurais que serão utilizadas, como, por exemplo, a arquitetura utilizada, assim como as métricas para o treinamento e análise.

Após isso, utilizaremos essas redes nas séries temporais escolhidas, comparando os desempenhos obtidos entre si e entre outros métodos já descritos na literatura.

Por fim, compilaremos os resultados para o relatório final, de forma a conter uma discussão ampla e representativa dos ensaios realizados e das conclusões obtidas.

Referências

- [1] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [2] J. T. Connor, R. D. Martin, and L. E. Atlas, “Recurrent neural networks and robust time series prediction,” *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 240–254, 1994.
- [3] N. Fiedler-Ferrara and C. P. C. do Prado, *Caos: uma introdução*. Editora Blucher, 1994.
- [4] S. Haykin, *Neural networks and learning machines*, 3/E. Pearson Education India, 2010.
- [5] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain.,” *Psychological review*, vol. 65, no. 6, p. 386, 1958.
- [6] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O’Reilly Media, 2019.
- [7] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [8] R. R. de Faissol Attux, “Sobre **dinâmica caótica e convergência** em algoritmos de equalização autodidata,” dissertação (mestrado), Universidade Estadual de Campinas, Faculdade de Engenharia **Elétrica** e de Computação, Campinas, SP, 2001.
- [9] D. Ruelle and F. Takens, “On the nature of turbulence,” *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25*, vol. 12, pp. 1–44, 1971.